COVER SHEET

TITLE: Exploring Senataxin – Sen1p Homology

AUTHOR’S NAME: Elizabeth A. Malone

MAJOR: Neurobiology and Psychology

DEPARTMENT: ICBE and Psychology

MENTOR: Michael R. Culbertson

DEPARTMENT: Genetics

MENTOR(2):

DEPARTMENT(2):

YEAR:

(The following statement must be included if you want your paper included in the library’s electronic repository.)

The author hereby grants to University of Wisconsin-Madison the permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.
ABSTRACT

Exploring Senataxin – Sen1p Homology

ABSTRACT: Mutated SETX, a human gene encoding for the protein Senataxin, has been identified among patients with two different neurodegenerative disorders - ataxia-oculo apraxia 2 (AOA2) and amyotrophic lateral sclerosis 4 (ALS4) (1). A cross-organism comparison was performed to identify possible orthologs of this newly identified protein. Saccharomyces cerevisiae protein, Sen1p, was determined to have two domains of homology with Senataxin: a matching Superfamily I helicase domain and a portion of the N-terminus to which Rad2p, Rnt1p, and Rpb1p interaction sites have been isolated in yeast (2). We will explore the degree of conservation between the proteins by determining if Senataxin interacted with the same three protein candidates – Rad2p, Rnt1p, and Rpb1p – as the yeast Sen1p by using a yeast two-hybrid test. Furthermore, we will examine if Senataxin interacts with the human orthologs of these three proteins. With these experiments, we hope to confirm the homologous functions of the two proteins in their respective organisms, which would allow for a greater understanding of SETX mutants, such as people affected by AOA2 and ALS4.

Elizabeth A. Malone / Neurobiology and Psychology
Author Name/Major

Michael R. Culbertson / Genetics
Mentor Name/Department

5/16/2010
Date

[Signature]
Author Signature

[Signature]
Mentor Signature
Exploring Senataxin – Sen1p Homology

Elizabeth A. Malone
University of Wisconsin-Madison

Michael R. Culbertson, PhD
Karen Chinchilla and Jonathan Finkel PhD
Department of Genetics and Molecular Biology
UW-Madison Graduate School

ABSTRACT: Mutated SETX, a human gene encoding for the protein Senataxin, has been identified among patients with two different neurodegenerative disorders - ataxia-ocular apraxia 2 (AOA2) and amyotrophic lateral sclerosis 4 (ALS4) (1). A cross-organism comparison was performed to identify possible orthologs of this newly identified protein. Saccharomyces cerevisiae protein, Sen1p, was determined to have two domains of homology with Senataxin: a matching Superfamily I helicase domain and a portion of the N-terminus to which Rad2p, Rnt1p, and Rpb1p interaction sites have been isolated in yeast (2). We will explore the degree of conservation between the proteins by determining if Senataxin interacted with the same three protein candidates – Rad2p, Rnt1p, and Rpb1p – as the yeast Sen1p by using a yeast two-hybrid test. Furthermore, we will examine if Senataxin interacts with the human orthologs of these three proteins. With these experiments, we hope to confirm the homologous functions of the two proteins in their respective organisms, which would allow for a greater understanding of SETX mutants, such as people affected by AOA2 and ALS4.
Introduction

A recent study of the human neuromotor disorders ataxia-ocular apraxia 2 (AOA2) and amyotrophic lateral sclerosis 4 (ALS4) uncovered mutations in a single gene, SETX (1). These mutations lead to inappropriate functioning of its protein product, Senataxin. Patients with these disorders typically develop peripheral neuropathy, with the phenotype of muscle atrophy, which greatly disrupts their mobility (3). Senataxin was found to be in highest concentration in the brain: specifically the cerebellum, hippocampus, and olfactory bulb in a rat model(4). At a subcellular level, the protein appeared to localize to the nucleoplasm (5).

Comparative analysis across organisms revealed that SEN1, a gene in the yeast Saccharomyces cerevisiae, is an ortholog of SETX (1). S. cerevisiae (baker's yeast) is a single-celled, eukaryotic organism that has served as a prominent organism for genetic research for years. Many characteristics make yeast a good model for research, such as its short generation time and simple life cycle, and because many biological functions in yeast have been well conserved in higher eukaryotes. Discoveries in yeast have contributed greatly to our understanding of more complex organisms including humans. Like Senataxin, the yeast protein product of SEN1, Sen1p, localizes to the nucleus where it functions as a DNA/RNA helicase (5,6). Sen1p can be divided into two segments. The N-terminal segment (amino acids 1-975) is important but not essential for viability since deletion of the entire segment results in slow growth. The C-terminal segment (amino acids 976-2231), which contains the DNA/RNA helicase active center and a nuclear localization signal (NLS), is essential for cell viability (6).

Human Senataxin shares significant sequence similarity with both the N- and C-terminal segments of yeast Sen1p (fig. 1). Using a global and candidate-specific yeast two-hybrid screen
to detect protein-protein interactions, it was found that the N-terminal 565 amino acids of Sen1p interacts with three different yeast proteins that are important to proper cellular function: Rad2p, an endonuclease that excises damaged DNA; Rnt1p, a protein involved in the 3’ end processing of non-protein encoding RNAs; and Rpb1p, the largest subunit of RNA polymerase II (2). The identification of protein-protein interactions provide important insights into how proteins function and which cellular pathways they are involved in. Mutations in Sen1p have been identified that disrupt its ability to interact with Rad2p, Rnt1p, and Rpb1p, resulting in altered cell morphology, UV-sensitivity, RNA processing, and transcription (2).

Mutations to a cell’s DNA occur in a random fashion. These mutations can either be silent (resulting in the same amino acid sequence), nonsense (resulting in a premature stop codon, leading to either a truncated protein or degradation), or a missense mutation (resulting in a new amino acid). Because of major changes to protein length and stability, nonsense mutations normally lead to pathogenic phenotypes. Missense mutations can also cause adverse consequences to the cell, especially if the mutation occurs in a functionally important region of the protein. Pathogenic phenotypes from missense mutations can result from a local change in the charge, or polarity, along the protein that can disrupt proper folding and thus the proper 3-dimensional shape needed for lock-and-key type interactions with other proteins. By cataloging the SETX mutations found in AOA2 and ALS4 patients, a density map of pathogenic missense mutations was created (fig. 2). Identifying the areas of Senataxin with a high density of
mutations can be helpful in locating functionally important domains. For example, figure two shows numerous mutations between amino acids 1931 and 2456, the region identified as the Helicase domain (I). A second region of high-density mutations lies in the N-terminal region of Senataxin. The evidence from the mutation density map, along with the sequence homology to the protein-interaction domain in the N-terminus of Sen1p, indicates a possible role for the N-terminus of Senataxin as an important site for protein-protein interactions.

![Mutation Density Map](image)

**Figure 2. Senataxin Mutation Density Map.** Missense mutations displayed in AOA2 and ALS4 patients are shown along the Senataxin protein at their relative locations. (Top) The protein was divided into 100 amino acid bins, and the number of missense mutations per bin was counted.

The goal of my research is to determine whether human Senataxin engages in protein-protein interactions similarly to yeast Sen1p. In this study we begin to compare the cellular
functions of Senataxin by using a yeast two-hybrid assay to test whether Senataxin can interact with the same three Sen1p-interacting proteins: Rad2p, Rnt1p, and Rpb1p. If we find Senataxin and Sen1p interact with the same candidate proteins, this would indicate, more so than just sequence comparison, their homology. Furthermore, if Senataxin-Sen1p homology can be verified, the understanding of the role of SEN1 in the S. cerevisiae would provide relevant insight into the general cellular functions of Senataxin and the consequences of SETX mutants, such as those associated with human genetic disorders.

The yeast two-hybrid test is used to detect possible protein-protein interactions in vivo. The gene of interest (bait) and the interaction candidates (prey) are placed into GAL4 binding domain (GBD) and GAL4 activating domain (GAD) vectors, respectively (7). GAL4 has traditionally been used because it has a transcription factor that can easily be divided into two distinct domains (binding and activating) without disrupting its ability to initiate transcription once reunited. Once the vectors are inside the yeast cell, transcription begins upstream of the GBD/GAD and continues through the inserted segment, creating a hybrid protein (Fig. 3). If the bait protein and prey candidate do interact, the binding domain and activation domain of GAL4 are brought together, resulting in the initiation of transcription of the downstream reporter gene. To test whether or not an interaction has occurred, the yeast cell can be placed on media lacking the essential nutrient coded for by the reporter gene. Viability would indicate that the missing nutrient is being produced and thus, that a protein-protein interaction is occurring.

Four segments of Senataxin were each cloned into the GBD vectors, while Rad2p, Rnt1p, and Rpb1p were cloned into GAD vectors (7). Each Senataxin construct, along with each of the prey vectors, is then introduced into the PJ694A yeast strain. The cells are grown on specific
media to select for cells able to activate the reporter gene, indicating an interaction between the two proteins.

![Diagram](image)

**Figure 3. Two-Hybrid Analysis of Senataxin Bait and Candidate Prey.** After transformation into the yeast cell, the bait (in blue) and prey (in green) vectors are translated into hybrid proteins, fused to the binding domain (red) or activating domain (orange), respectively. The BD binds to the promoter on the yeast chromosome. If the AD is then brought into close proximity with the BD by bait-prey interaction, transcription of the reporter gene, *HIS3* (pink), is initiated.

**Materials & Methods**

**Cloning of DNA Segments into Vectors**

Full-length cDNA of *SETX* was generously provided by Professor Craig Bennett at the University of Washington. Primers corresponding to different regions of *SETX* were designed and used to create smaller segments of *SETX* using PCR. These fragments were named DU1 (aa 1-702), EM2 (aa 630-1284), EM3 (aa 1250-1933), and EM4 (aa 1900-2677[end]). These four approximately equal segments of *SETX* were used as bait and were cloned into pGBD-C1 and pGBDU-C1 vectors (7). These plasmids both contain the exact same *GAL4*-binding domain segment, and differ only by the selective marker: tryptophan for GBD and uracil for GBDU. To confirm the presence of each *SETX* piece in the pGBD/pGBDU vectors, we sequenced the constructs using forward and reverse primers that recognized sequences upstream and
downstream of the insert, respectively. The three candidate prey plasmids, GAD-RAD2, GAD-RNT1, and GAD-RPBI, were identified and isolated using the P. James C-1 library in a yeast two-hybrid screen (2). The GAD plasmid contains leucine as the selective marker.

Candidate-Specific Yeast Two-Hybrid Analysis

The bait and prey constructs were transformed into *S. Cerevisiae* strain PJ69-4a using the lithium acetate transformation method (7). This strain has the reporter genes *HIS3, ADE2*, and *lacZ* under the control of the *GAL4* promoter. The presence of the plasmids in the cells was confirmed by plating cells on media lacking leucine and tryptophan or uracil. Interactions were detected by plating cells on media lacking adenine or histidine and assaying for growth. Serial dilutions of each transformant were used to monitor the rate of cell growth (2). The strength of the interactions on the histidine plates was evaluated by using increasing concentrations of the drug 3-aminotriazole (3AT) - 0.0 mM, 1.0 mM, 2.5 mM, and 5.0 mM – where an interaction occurring at 5.0 mM is the strongest.

![Figure 4. Two hybrid interactions with SETX fragments. Control plates are shown for all transformants indicating the presence of both vectors by using media lacking their selective markers. Transformants are also shown on media lacking histidine with 1 mM of 3AT. Growth on this plate would indicate activation of the reporter gene HIS3 and thus a protein-protein interaction. A) Cells were transformed with DU1 (aa 1-702) in a GBD plasmid and each candidate in a GAD plasmid. Interaction between Senataxin and Rnt1p is observed. B) EM2-GBDU interacted with Rad2p and GAD. C) and D) No interactions were observed between EM3-GBD or EM4-GBD and the GAD-prey proteins.](image-url)
Results

Detection of a protein-protein interaction involves growth on a plate lacking the nutrient that is expressed by one of the reporter genes used (adenine or histidine). This is due to the activation of the ADE2 or HIS3 gene via the bringing together of the GAL4-activating domain and GAL4-binding domain.

Yeast transformants containing the DU1-GBD/GBDU vector and the prey-GAD vectors were spotted on media lacking leucine and tryptophan or uracil (Materials and Methods) to select for both plasmids (Fig. 4). Transformants containing DU1-GBD and each prey vector were then spotted on media lacking leucine, tryptophan, and histidine (Materials and Methods). Senataxin, aa 1-702 (DU1-GBD), and Rnt1p were able to activate the HIS3 reporter, indicating a protein-protein interaction. The strength of the interaction was shown by growth on media containing 1 mM 3AT and, to a lesser extent, on 2.5 mM 3AT media. Adenine, another, more stringent reporter gene, was also used to assay growth. Spotting on plates lacking leucine, tryptophan, and adenine resulted in no significant growth (data not shown). These experiments were repeated using DU1-GBDU and the results were similar to those seen for DU1-GBD (data not shown).

Protein-protein interactions between EM2-SETX (aa 630-1284) and the three GAD-fusion candidates were assayed in the same manner as in figure 3a, except pGBDU was used. EM2-Rad2p interaction was detected as evidenced by growth on media lacking histidine. The strength of the interaction was determined plating cells on media containing 1 mM 3AT concentration (Fig. 3b). The majority also showed interactions on 2.5 mM 3AT plates (data not shown). Surprisingly, an interaction between EM2 and GAD was also detected and shown to be strong by growth on plates containing 3AT. Because of the strong interaction with our negative control, GAD-empty vector, no conclusive data can be drawn from this test. Media lacking leucine, uracil, and adenine showed no growth. The two-hybrid experiments performed with the
EM2-GBD plasmid showed similar results (data not shown). Due to the unusual nature of these results, new transformations were performed several times with both the pGBD and pGBDU to verify the previous results. These assays were not significantly different from the first results and demonstrated the same interactions.

EM3-SETX (aa 1250-1933) and EM4-SETX (aa 1900-2677), did not interact with any of the candidate proteins (Fig. 3c and 3d). No growth was detected on media lacking adenine or histidine. The same findings were observed with EM3-GBDU and EM4-GBDU (data not shown).

**Discussion**

In an attempt to look for similarities in function between Sen1p and Senataxin, a yeast two-hybrid test was used to assay for interactions between Senataxin fragments and the yeast Sen1p-interacting proteins. Results from the DU1-SETX segment showed an interaction with Rnt1p. The interaction appeared similar to Sen1p-Rnt1p, being detectable on media lacking histidine containing 1 mM 3AT. In the yeast cell, Rnt1p is involved in the 3’end processing of RNA’s. Interaction of Rnt1p with Senataxin may indicate that, like Sen1p, Senataxin is involved in the 3’ end processing of both protein-encoding and non protein-encoding RNAs. The experiments using EM2 as the bait produced inconclusive data. Since the vector alone interacted with this piece of Senataxin, we must be very critical of any other positive result, such as the one detected with Rad2p. An interaction with GAD, the plasmid without any inserted prey protein, is normally indicative that your bait protein is self-activating. However, since a self-activating protein by definition would require no outside interaction to initiate the transcription of the reporter gene (*HIS3*), one should find activation of all transformants that contain that bait vector, regardless which, or if any prey is present. Since no growth was recorded for either the
Rpb1p-GAD or Rnt1p-GAD transformants, we can conclude that EM2-SETX is not a traditional self-activator. One possible, yet unlikely explanation for these results is that EM2-SETX has self-activating capabilities that only express conditionally. This explanation, however, does not adequately account for the interaction with Rad2p and with none of the other candidate proteins, assuming they were exposed to the same conditions.

We proposed another theory to explain these unusual results. It is possible that the piece of Senataxin contained within EM2 has a false positive interaction with part of the translated GAL4 activating-domain (GAD). However, once the prey protein is fused to the AD, creating the hybrid, the accessibility to the AD is altered due to folding, no longer allowing an interaction to occur. This could be why no interaction is seen for Rnt1p or Rpb1p even though the same AD protein piece is present. On the other hand, Rad2p-GAD transformants demonstrated interaction with Senataxin. With Rad2p, we cannot accurately distinguish between a true Senataxin-Rad2p interaction, and a Senataxin-AD interaction. If the Rad2p-EM2 interaction is indeed real, Senataxin could be involved in nucleotide excision repair in the human cell. If this result were confirmed, it would be particularly interesting in terms of its applicability; it has been shown that several autosomal recessive ataxias in humans, including AOA2, are categorized by a defective response to DNA damage (5). If Rad2p interacts with Senataxin, it could provide a useful starting point for investigating the role of Senataxin in DNA repair. More research in this area could also lead to a specific, molecular explanation of the broader phenotypic ramifications of SETX mutants.

Both of these theories mentioned above could explain the results we obtained, yet more testing is certainly required to validate either. A third explanation is that in arbitrarily cutting Senataxin into quarters, we disrupted an important domain and are therefore getting inconclusive
results. This, possibly in combination with the second theory proposed, seems most feasible to us. In future experiments, we hope to clone the whole first half of SETX into the pGBD and pGBDU vectors and perform new two-hybrid transformations and to test with the same candidate proteins. If the results are still inconclusive after this point, the next step would be to switch the bait and prey vectors; Rad2p, Rnt1p, and Rpb1p would be placed into the BD vectors and the SETX piece would go into the AD, theoretically preventing any self-activation or interaction with the AD from disturbing the results.

The last two segments of Senataxin, EM3 and EM4, interacted with none of the candidate proteins. These results were expected when considering the homology between SETX and SEN1; EM3 and EM4 together roughly correlate to the C-Terminal domain of Sen1p, which was similarly found to not interact with Rad2p, Rnt1p, and Rpb1p (2).

Because the interaction candidates are S. Cerevisae proteins, we cannot yet say that their interactions with the human Senataxin are functionally relevant. Even though our results don’t fully show Sen1p and Senataxin to be homologous proteins, they do give us relevant information about the degree of conservation of their functional domains. Like Sen1p, the C-terminus of Senataxin seems unimportant for binding Rad2p, Rnt1p, and Rpb1p proteins, while the n-terminus is necessary. The evolutionary conservation of the candidate proteins that were used, leaving them with orthologs in human cells, enhances the applicability of our findings. Rpb1p is the most similar to its human ortholog. The C-terminal domain (CTD) of the protein, which has been identified as the region responsible for binding Sen1p, is identical in the orthologous portion of RNA Polymerase II. Rad2p also shares an ortholog in humans called XPG, which has been shown to have similar structural specificity and function in the different organisms.

Finally, Rnt1p has two human proteins with which it shares homology: Dicer and
Drosier. Further direction of this study could involve testing for interaction of Senataxin with the human orthologs just mentioned. If interactions are identified, this would provide strong evidence that function is conserved. Also, a map of the locations of these interactions on Senataxin could be compared to the mutation density map (fig. 2) to determine if missense mutations that disrupt important protein-protein interactions could be causing the disease phenotypes of AOA2 and ALS4. If this approach is successful, it may be possible to predict the underlying biochemical defects in patients carrying mutations at known locations in the SETX gene.

References


(3) T Asaka et al. 2006. Autosomal recessive ataxia with peripheral neuropathy and elevated AFP: novel mutations in SETX. Neuro. May 23(10): 1580-1581


